

Search Notes: 10/572, 638

DATE: Thursday, December 16, 2010

Set Side	Name by	Query	Hit Count	Set Result	Name Grid	Set Name Set
----------	---------	-------	-----------	------------	-----------	--------------

Interference Searches

DB=PGPB,USPT,UPAD; PLUR=YES; OP=AND

L16	L15 not (L3 or L6 or L10 or L8)	34	L16	L16
L15	L14 and CON-S	45	L15	L15
L14	L13 and (env\$)	45	L14	L14
L13	L12 and (CON-S\$)	48	L13	L13
L12	L11 and (consensus or ancestral)	25969	L12	L12
L11	(HIV\$ or (human immunodeficiency virus))	104532	L11	L11
L10	L9 not (L3 or L6)	7	L10	L10
L9	hahn.in. and beatrice.in.	15	L9	L9
L8	L7 not (L3 or L6)	8	L8	L8
L7	korber.in. and bette.in.	10	L7	L7
L6	L5 not L3	6	L6	L6
L5	L4 and (consensus or ancestral)	8	L5	L5
L4	gao.in. and feng.in.	242	L4	L4
L3	L2 and CON-S\$	10	L3	L3
L2	L1 and (consensus or ancestral)	35	L2	L2
L1	haynes.in. and barton.in.	67	L1	L1

END OF SEARCH HISTORY

FILE 'WPIDS' ENTERED AT 11:53:27 ON 16 DEC 2010

	E HAYNES B F/IN
L1	65 S E3
	E HAYNES B/IN
L2	38 S E3
L3	90 S L1 OR L2
L4	50 S L3 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L5	5 S L4 AND (CONSENSUS OR ANCESTRAL)
	E GAO F/IN
L6	1996 S E3
L7	10 S L6 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L8	7 S L7 NOT L5
L9	0 S L8 AND (CONSENSUS OR ANCESTRAL)
	E KORBER B T/IN
L10	9 S E2 OR E3
	E HAHN B H/IN
L11	15 S E3
	E HAHN B/IN
L12	155 S E3
L13	166 S L11 OR L12

L14 10 S L13 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L15 35091 S (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L16 370 S L15 AND (CONSENSUS OR ANCESTRAL)
L17 71 S L16 AND ENV?
L18 1 S L17 AND (CON-S)

FILE 'MEDLINE' ENTERED AT 11:56:38 ON 16 DEC 2010
E HAYNES B/AU

L19 56 S E3
E HAYNES B F/AU

L20 260 S E3

L21 316 S L19 OR L20

L22 53 S L21 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L23 0 S L22 AND (CONSENSUS OR ANCESTRAL)
L24 0 S L22 AND (CON-S)
L25 22 S L22 AND (ENV?)
E GAO F/AU

L26 306 S E3
E GAO FENG/AU

L27 699 S E3

L28 1005 S L26 OR L27

L29 80 S L28 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L30 15 S L29 AND (CONSENSUS OR ANCESTRAL)
E KORBER B T/AU

L31 141 S E3-E6

L32 132 S L31 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L33 65 S L32 AND ENV?
L34 2 S L33 AND (CON-S)
E HAHN B H/AU

L35 208 S E3
E HAHN BEATRICE H/AU

L36 91 S E3
E HAHN B/AU

L37 88 S E3

L38 387 S L35 OR L36 OR L37

L39 175 S L38 AND (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L40 23 S L39 AND (CONSENSUS OR ANCESTRAL)
L41 17 S L40 AND ENV?
L42 4 S L41 NOT L30

L43 222604 S (HIV? OR HUMAN IMMUNODEFICIENCY VIRUS)
L44 9 S L43 AND (CON-S)

L45 2017 S L43 AND (CONSENSUS OR ANCESTRAL)
L46 502 S L45 AND ENV?
L47 70 S L46 AND (FUSION OR CLEAVAGE)
L48 48 S L47 AND PY<2004
L49 40 S L48 AND (CONSENSUS/AB OR ANCESTRAL/AB)

CON-S is consensus group M polypeptide obtained from the Los Alamos 2000 sequence listing; it contains a deletion of the fusion domain (~ aa 475-505) and a C-terminal stop codon at ~aa 645.

Consensus env gene is constructed by generating consensus sequences of env genes for each subtype of a particular group (e.g., group M comprises subtypes A-D, F-H, J, and K) from the Los Alamos HIV sequence database. A consensus sequence of all subtype consensus can be generated to avoid heavily sequenced subtypes. [p. 38]

CON-S was derived from the year 2000 database group M env sequences. [p. 40]

Con-S gp140 induced robust neutralization responses in multiple group M subtypes (e.g., B and C).

CON-S was obtained by aligning the consensus env sequences of group M subtypes (A-D, F-H, J, and K) from the 2003 Los Alamos sequence listing to obtain a parent sequence. The hypervariable regions (excluding V3) were generated through hand alignments of potential N-linked glycosylation sites and cysteine residues. The hypervariable loops tended to comprise a shorter range of lengths as compared to natural strains. The CON-S env gene was generated by converting amino acid sequences of CON-S to nucleotide sequences employing the codon usage of highly expressed human genes. HIV-1 gp140 Envs with deletion of the cleavage (C) site and fusion (F) domain were designated gp140ΔCF.

Fig. 29A

CON_OF_CON-S-2003 (829 a.a.)

MRVNGIÖRNCOHLRWRGILIFGMLIICSAEENLWTVYGVGVVKEARNTTLCASDAKAYDTEVHNWATHACVPTDPNPQEIVL
ENVTEFNFMWKNNNVEQNHEDIISLMDQSLKECVVLTLPELCVTLNCTDVNATNNNTNNNEEIKNCSFNITTEIRDKKKVYALFYKL
DVVPIDONNSYRLINCNTSAITQACPKVSEFIPINHYCAPAGFAILKCNKNDKKFNGTCPKNVSTVQCTHGEKPVVSTOLLNGSL
AEEETIIRSEMINITNNAKTIIIVQNLNESEVINECTRPNNNTRKSIRIIGPQAFYATGDIIGDIRQAHCNISRKTGNKTLQOVAKKLRE
HFNKTIIIFNPSSGGDLEITTHSFnNCGGEFYCNTSELFNSTWNGTNNTTILPCRKIGIINMMWQGVQGAMYAPPIEGKIRCTSNT
GILLTRDGNNNTETFRPGGGDWRONWRSELYKVVKIEPLGVAPTKAKRVRVREKRAVGIGAVFLGFLGAAGESTNGAASITL
TVQCARQLLSGIVQQQSNLRLRAIEAQOHILLQITVWGIQQLQARVLAVEREYLKDQQLGIGCNSGKRLICITNNFWNNSWSNKSQDST
WDNNMTWMEWOKEINNYTDIYSLIEESQMQQKEKNEQELLALDKWASLNNWFDDITNWLWYIKIFIMIVGGLIGLRLIVFAVLIVNR
VRQGYSPLSFTLIPNFRGDPDRPEGIEEEGGEQQRDRSIRLVLNGFLALAWDDLSLCLFSYHRLRLDILIJARTVELLGRGWERA
LKYLNLLQYWQGQELKNSAISLLDTAIAVAEGTDRVIEVQVQRVCRAILNIPRRRIQGFERALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 29B.

CON-S-2003 140CE.pep (620 a.a.)

Nick name: 006

MRVNGIQRNCQHLWRWGILIFGMLIICSAEHNVTYYGVPVKEANTTLFCASDAKAYDTEVHNWATHACVPIDPNPQEIWL
ENVTENEWMNNKVNNEQHMEDISLWQSDQVCKVLPFLTCVLCNTDVTWNTNNTNEEINCKNSNITTEIDRKKVWYALFYK
DPPVDDNNNSYRLINCNSTSATIQCACPQKFSVEPIHCPQAGAPFLCKNDKFCNGTQPGCKNVTSTCQTHGKIEFVPTQQLLNGSL
AEEELIISRENTIINNAAITIQLVNEEINCPTRPNNTNTRKSIRIIGPCQAFYATGDLIDIGRNCNSIRTSRKWNKLTQVOVAKLRE
HFNKTIITFNPSSGDELITTHSFCNGGCFYQNTCLESFNSTWNGTNTTILTPCRKQIINNMQVGQGAMYAPIZEGKRCITSN
GLLRLTRDGGNNNTETFRBGGCDMRDNWRSLEYKVKVKEPLGVAPTKAKTLTVQARQLLSGIVQGSNLLRAIEAQQLRQLTV
WGKLOQARVLAVLERYKLIDQDGLGNGCSCGRLICITNWFMISSWSNKSQDEIDWNTWMEIDKIEINNYTDIISLIEESQNQZK
NECELLADLWNSLWNPDITLNWL

*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

Fig. 29C

CODON-OPTIMIZED CON-S-2003 140CF, seq (1891 nt)

Nick name : 006

The invention describes an isolated protein (I) selected from 106 fully defined 500-866 amino acid sequences given in the specification. Also described are: a nucleic acid comprising: a nucleotide sequence encoding CON6 HIV gp160 protein, subtype C ancestral HIV envelope protein, subtype C consensus HIV envelope protein, subtype C consensus HIV gag protein, subtype C consensus HIV nef protein, Group M consensus HIV envelope protein, subtype A consensus HIV envelope protein, Group M consensus HIV gag protein, Group M consensus HIV pol protein, Group M consensus HIV nef protein, subtype C consensus HIV pol protein, subtype B consensus HIV gag protein, or subtype B consensus HIV envelope protein, where the nucleotide sequence comprises codons optimized for expression in human cells; a nucleotide sequence encoding (I); or a nucleotide sequence selected from 89 fully defined 633-2607 bp sequences given in the specification; a vector comprising the nucleic acid of (I); a composition comprising at least one protein or nucleic acid above and a carrier; and inducing an immune response in a mammal. The protein is a consensus or ancestral immunogen useful for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or that induces a T cell response. This is the amino acid sequence of HIV CON-3 env protein.

xx

so Sequence 829 AA;

Query Match 99.2%; Score 3338.5; DB 2; Length 829;
Best Local Similarity 94.9%;
Matches 620; Conservative 0; Mismatches 0; Indels 33; Gaps 1;

Qy 1 MRVVMGIQRNCQHLWRNGILIFGMLIICSAEENLWVTYYGVPVWKEANTTLFCASAKAY 60
|||
Db 1 MRVVMGIQRNCQHLWRNGILIFGMLIICSAEENLWVTYYGVPVWKEANTTLFCASAKAY 60

Qy 61 DTEVHNW@ATHACVPTDPNPQEIVLENVTENFNM@KNNMVEQMHEDEIISLWDQSLKPCVK 120
|||
Db 61 DTEVHNW@ATHACVPTDPNPQEIVLENVTENFNM@KNNMVEQMHEDEIISLWDQSLKPCVK 120

Qy 121 LTPLCVTLNCTD@NATNNTTNEEIKNCFSNITTEIRDKKKVVYALFYKLDVVPIDDNN 180
|||
Db 121 LTPLCVTLNCTD@NATNNTTNEEIKNCFSNITTEIRDKKKVVYALFYKLDVVPIDDNN 180

Qy 181 YRLINCNTSAITQACPKVSFEPPIHYCAPAGFAILKCNDFNGTGPCKNVSTVQCTHG 240
|||
Db 181 YRLINCNTSAITQACPKVSFEPPIHYCAPAGFAILKCNDFNGTGPCKNVSTVQCTHG 240

Qy 241 IKPVVSTQLLNLNGSLAEEEEEIRGENITNNAKTIIIVQLNESVEINCTRPNNNTRKSIRIG 300
|||
Db 241 IKPVVSTQLLNLNGSLAEEEEEIRGENITNNAKTIIIVQLNESVEINCTRPNNNTRKSIRIG 300

Qy 301 PGQAFYATGDIIGDIRQAHCNISRTKWNKTLQQVAKKLREHFNKTIIIFNPSSGGDLEITT 360
|||
Db 301 PGQAFYATGDIIGDIRQAHCNISRTKWNKTLQQVAKKLREHFNKTIIIFNPSSGGDLEITT 360

Qy 361 HCFNCGGEFFYCNTSELFNSTWNGTNNTITLPCRIKQIINMMWQGVGQAMYAPPIEGKIRC 420
|||
Db 361 HCFNCGGEFFYCNTSELFNSTWNGTNNTITLPCRIKQIINMMWQGVGQAMYAPPIEGKIRC 420

Qy 421 TSNITGLLLTDGGNNNTETFPRPGGDMRDNWRSLEYKVKVVKIEPLGVAPTKAK---- 475
|||
Db 421 TSNITGLLLTDGGNNNTETFPRPGGDMRDNWRSLEYKVKVVKIEPLGVAPTKAK**RRVVE** 480

Qy 476 -----TLTVQARQLLSGIVQQGSNLLRAIEAQQLLQ 507
|||
Db 481 REKRAVGIGAVFLGFLGAAGSTMGRASITLTVQARQLLSGIVQQGSNLLRAIEAQQLLQ 540

Qy 508 LTVWGIKQLQARVLAVERYLKDQQLLGIGWCGSGKLICCTTIVPWNSSSWNKSQDEIWDNMT 567
|||
Db 541 LTVWGIKQLQARVLAVERYLKDQQLLGIGWCGSGKLICCTTIVPWNSSSWNKSQDEIWDNMT 600

Qy 568 WMEWDKEINNYTDIYSLIEESQNQQEKNEQELLALDKWASLWNWFDTNWLNW 620
|||
Db 601 WMEWDKEINNYTDIYSLIEESQNQQEKNEQELLALDKWASLWNWFDTNWLNW 653